



Data Centre NetCDF Implementation Pilot

Peter Miu



EUMETSAT User Conference

Oslo, Sep. 2011

Splinter Session, Facilitating Data Access and Utilisation

What is this Pilot all about ?

The Need for the Data Centre to have a Common Delivery Format.

- Experience has shown that the User Community who orders Data Centre Archive products in a 'data' format need support in using these formats.
- **Formats such as BUFR and GRIB are excellent for NRT- dissemination but are difficult to work with (especially for the Non-Meteorological User Community) as the data is encoded in order to compress the size of the product.**
- Based on these inputs, the EUMETSAT Format Advisory Group proposed to make NetCDF the common delivery format for all products in the Data Centre Archive, which was approved by the EUMETSAT delegate body.

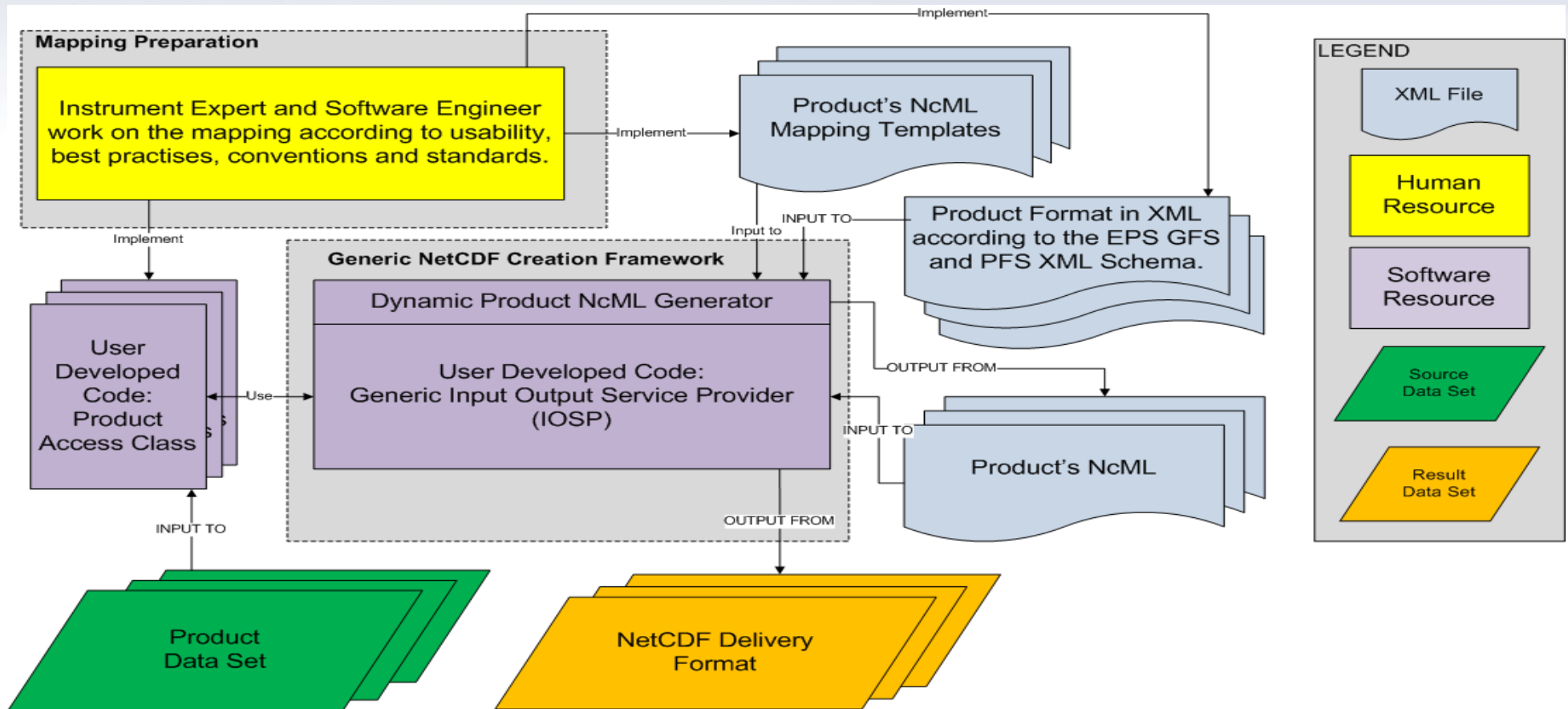
Why a Pilot ?

There are a lot of products (ca. 70) in the Data Centre Archive. The idea of the pilot is to investigate:

- The scale of the development; what resources are needed, what needs to be considered in the development (conventions, etc.) and identify a realistic time estimate for the full implementation.
- **By preparing a NetCDF pilot, complexity is reduced and a concept is established which is needed for the full implementation.**
- The pilot will focus on development of NetCDF formats for the ASCAT level 1 and level 2 products. These products were selected as they are popular with the user community and experts (OSI SAF) can support the validation of the products created.

What has been Implemented ?

The following Generic NetCDF Creation Framework has been implemented to support the creation of the NetCDF formats.



Attributes of the Created NetCDF Formats

JAVA NetCDF4 libraries are used to create NetCDF files that following the **Classic Data Model** to ensure compatibility with existing NetCDF tools. **Climate and Forecast Conventions** have been applied where applicable. **WMO File Naming Conventions** are used to describe the contents of the NetCDF file. Best Practises applied such as avoiding floats and doubles, replace with **shorts, integers and longs associated with a scaling factor**; helps compression.

- NetCDF ASCAT GDS L1b Sigma0 (ASCSZO1B - 25km) Native Size ~6MB
- **W_XX-EUMETSAT-Darmstadt,SURFACE+SATELLITE,METOPA+ASCAT_C_EUMP_20110620020002_24214_eps_o_250_I1.nc ~6MB**
- NetCDF ASCAT GDS L1b Sigma0 (ASCSZR1B - 12.5km) Native Size ~25MB
- **W_XX-EUMETSAT-Darmstadt,SURFACE+SATELLITE,METOPA+ASCAT_C_EUMP_20110620020002_24214_eps_o_125_I1.nc ~25MB**
- NetCDF ASCAT L2 Soil Moisture (ASCSSMO02 - 25km) Native Size ~9MB
- **W_XX-EUMETSAT-Darmstadt,SURFACE+SATELLITE,METOPA+ASCAT_C_EUMP_20110620020002_24214_eps_o_250_ssm_I2.nc ~4MB**
- NetCDF ASCAT L2 Soil Moisture (ASCSSMR02 - 12.5km) Native Size ~35MB
- **W_XX-EUMETSAT-Darmstadt,SURFACE+SATELLITE,METOPA+ASCAT_C_EUMP_20110620020002_24214_eps_o_125_ssm_I2.nc ~14MB**

Where can I get Example ASCAT NetCDF Files ?

- Demonstration Products are available from the EUMETSAT GSICS Server.

The screenshot shows a web browser window displaying the THREDDS catalog for the EUMETSAT GSICS server. The URL <http://gsics.eumetsat.int/thredds/catalog.html> is circled in red. The page lists various datasets in a table with columns for Dataset, Size, and Last Modified. The 'ASCAT DEMOSTRATION PRODUCTS' entry is highlighted with a magnifying glass.

Dataset	Size	Last Modified
GSICS Source Data		--
EUMETSAT/		--
CNES/		--
JMA/		--
GSICS Intermediate Data		--
EUMETSAT/		--
JMA/		--
GSICS Products		--
EUMETSAT/		--
JMA/		--
GRAS COLLABORATED DATA SETS		--
EUMETSAT/		--
GFZ/		--
ASCAT DEMOSTRATION PRODUCTS		--
EUMETSAT/		--

ASCAT DEMOSTRATION PRODUCTS

EUMETSAT Data and Product Server at www.eumetsat.int
THREDDS Data Server [Version 4.2.3 - 20110113.2322] Documentation

THREDDS: Data Set Meta-Data Information and Services

gsics.eumetsat.int/thredds/catalog/Level225Km/catalog.html?dataset=ascatLevel225Km/W_XX-EUMETSAT-Darmstadt,SURFACE%2BSATELLITE,METOPA%2BASCAT_C_EUMP_20110620135100_24221_eps_o_125_ssm_l2.nc

  **EUMETSAT Data and Product Server**
THREDDS Data Server

Catalog <http://gsics.eumetsat.int/thredds/catalog/Level225Km/catalog.html>

Dataset: Level 2 25 Km/W_XX-EUMETSAT-Darmstadt,SURFACE+SATELLITE,METOPA+ASCAT_C_EUMP_20110620135100_24221_eps_o_125_ssm_l2.nc

Attributes of the File

- Data format: NetCDF
- Data size: 14.03 Mbytes
- Naming Authority: www.eumetsat.int
- ID: ascatLevel225Km/W_XX-EUMETSAT-Darmstadt,SURFACE+SATELLITE,METOPA+ASCAT_C_EUMP_20110620135100_24221_eps_o_125_ssm_l2.nc

Documentation:

- Summary: Level 2 25 Km
- [EUMETSAT Format Guide documentation](#)
- Rights: Freely available

OpenDAP service to examine the NetCDF File contents

Access:

1. OPENDAP: /thredds/dodsC/Level225Km/W_XX-EUMETSAT-Darmstadt,SURFACE+SATELLITE,METOPA+ASCAT_C_EUMP_20110620135100_24221_eps_o_125_ssm_l2.nc
2. HTTP Server: /thredds/fileServer/Level225Km/W_XX-EUMETSAT-Darmstadt,SURFACE+SATELLITE,METOPA+ASCAT_C_EUMP_20110620135100_24221_eps_o_125_ssm_l2.nc

Keywords:

- EUMETSAT
- NetCDF

Dates:

- 2011-08-31 15:12:47Z (modified)

Creators:

- EUMETSAT
 - email: ops@eumetsat.int
 - <http://www.eumetsat.int>

Publishers:

- EUMETSAT
 - email: ops@eumetsat.int
 - <http://www.eumetsat.int/>

Viewers:

- NetCDF-Java ToolsUI (webstart)

Documentation links

Download link


THREDDS: OpenDAP Client accessing a ASCAT NetCDF File on the GSICS server

- An example of a free tool available for examining the contents of the NetCDF files:

← → ↻ 🏠 gsics.eumetsat.int/thredds/dodsC/Level225Km/W_XX-EUMETSAT-Darmstadt,SURFACE+SATELLITE,METOPA+ASCAT_C_EUMP_20110620135100_24221_eps_o_125_ssm_i2.nc.html

OPeNDAP Dataset Access Form

Tested on Netscape 4.61 and Internet Explorer 5.00.

Action:  **Variable contents can be displayed and saved in ASCII or Binary form**

Data URL:

Global Attributes:

```
creator_name: "EUMETSAT"
creator_url: "http://www.eumetsat.int"
creator_email: "ops@eumetsat.int"
institution: "EUMETSAT"
license: "CopyRight EUMETSAT 2011"
```

Variables:

time: Array of 32 bit Integers [numRows = 0..3162]

numRows:

```
_FillValue: "-2147483647"
Valid_min: "0"
valid_max: "2147483647"
standard_name: "time"
long_name: "time"
```

cell_index: Array of 16 bit Integers [numRows = 0..3162][numCells = 0..81]

numRows: numCells:

```
_FillValue: "-32767"
Valid_min: "1"
valid_max: "42"
standard_name: "cell index"
long_name: "across track cell index"
```

swath_index: Array of 8 bit Bytes [numRows = 0..3162][numCells = 0..81]

numRows: numCells:

```
_FillValue: "127"
flag_values: "0b, 1b"
flag_meaning: "left, right"
standard_name: "swath index"
long_name: "left/right swath index"
```

lat: Array of 32 bit Integers [numRows = 0..3162][numCells = 0..81]

numRows: numCells:

```
_CoordinateAxisType: "Lat"
Standard_name: "latitude"
long_name: "latitude"
units: "degree north"
valid_min: "-90000000"
```


Free NetCDF Tools: ToolsUI (Unidata) – ASCAT L2

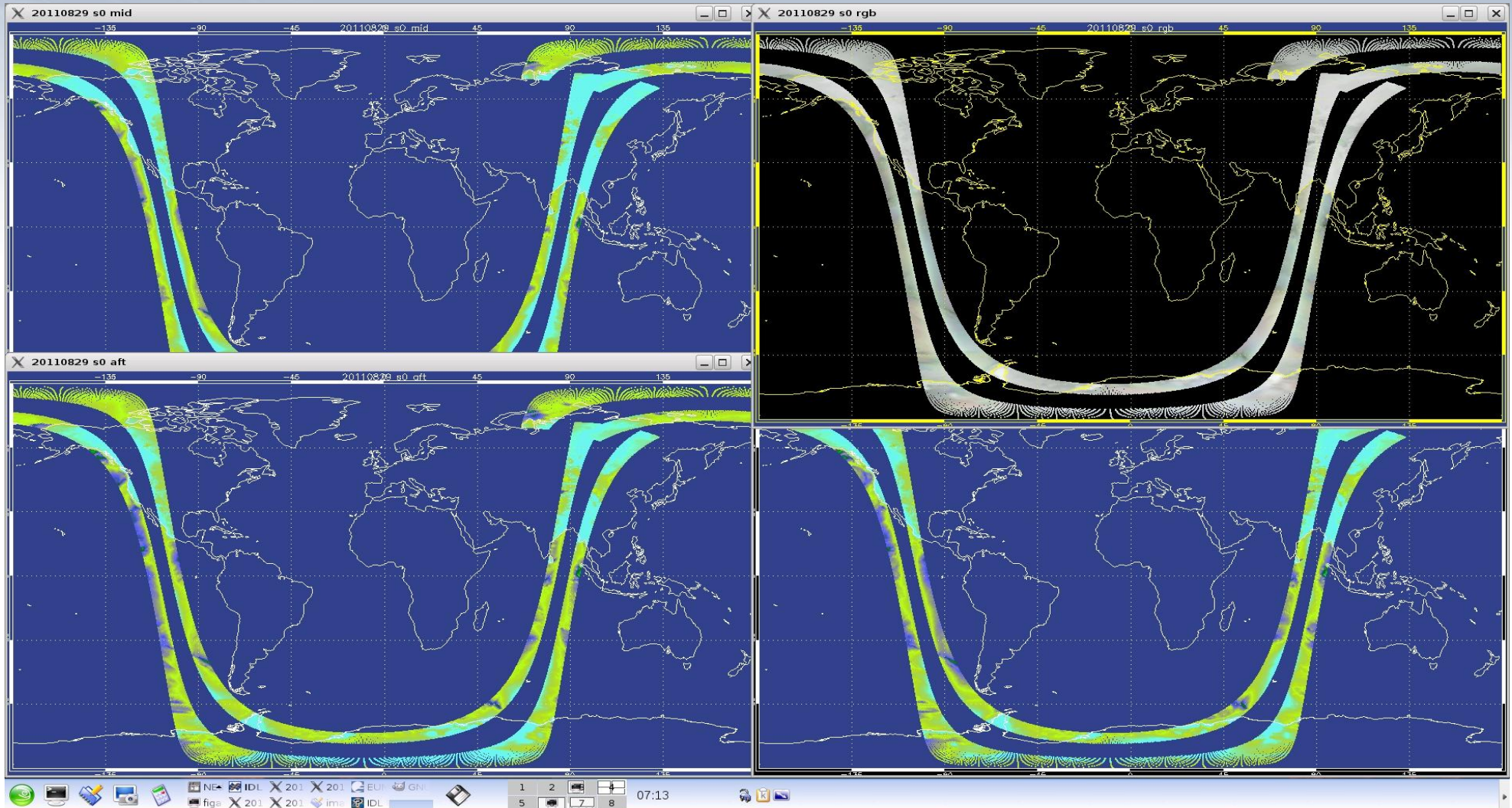
The screenshot displays the NetCDF Tools application interface. The main window shows a list of variables with columns for dataType, description, dimensions, group, name, shape, and units. A variable 'soil_moisture' is highlighted in blue. An arrow points from this variable to a text label: "All variables in the NetCDF File".

Below the variable list, another arrow points to a pop-up window titled "NCDump Variable Data". This window displays the metadata and data for the selected variable. The metadata includes:

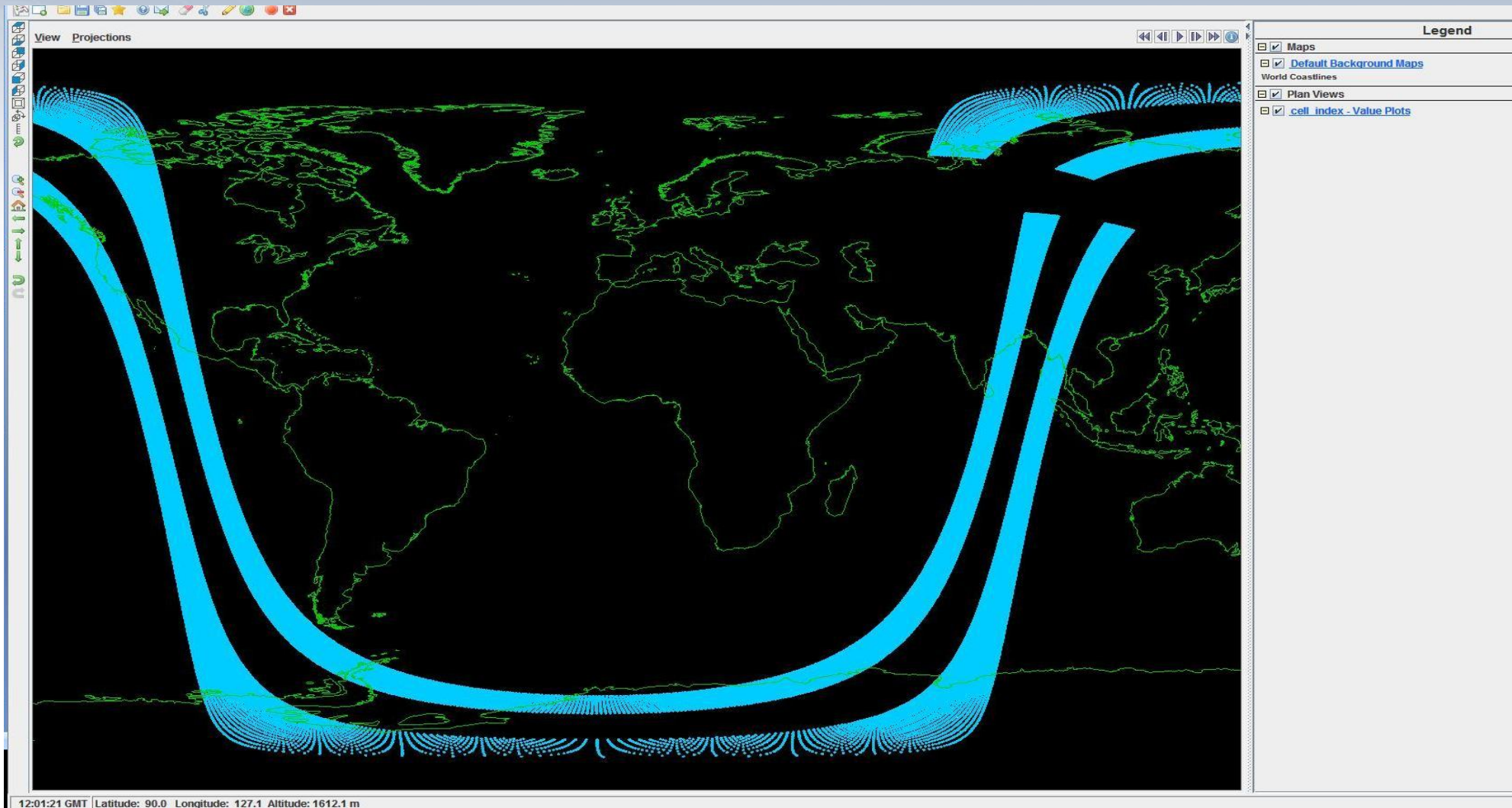
- Variable: soil_moisture(0:3258:1, 0:81:1)
- Dimensions: short soil_moisture (numRows=3259, numCells=82)
- FillValue: "-32767"
- Valid Range: "0" to "10000"
- Standard Name: "soil_moisture"
- Long Name: "Surface soil moisture"
- Units: "percentage"
- Quality Flag: "soil_moisture_error, soil_moisture_sensitivity, rainfall_flag, correction_flag, processing_flag"
- Coordinates: "lat lon"
- Scale Factor: "0.01f" (float)
- Comment: "Mean surface soil moisture (0 to 100%)"

The data section shows a large array of numerical values representing soil moisture percentages. A text label "Contents of a variable can be displayed" with an arrow points to this data section.

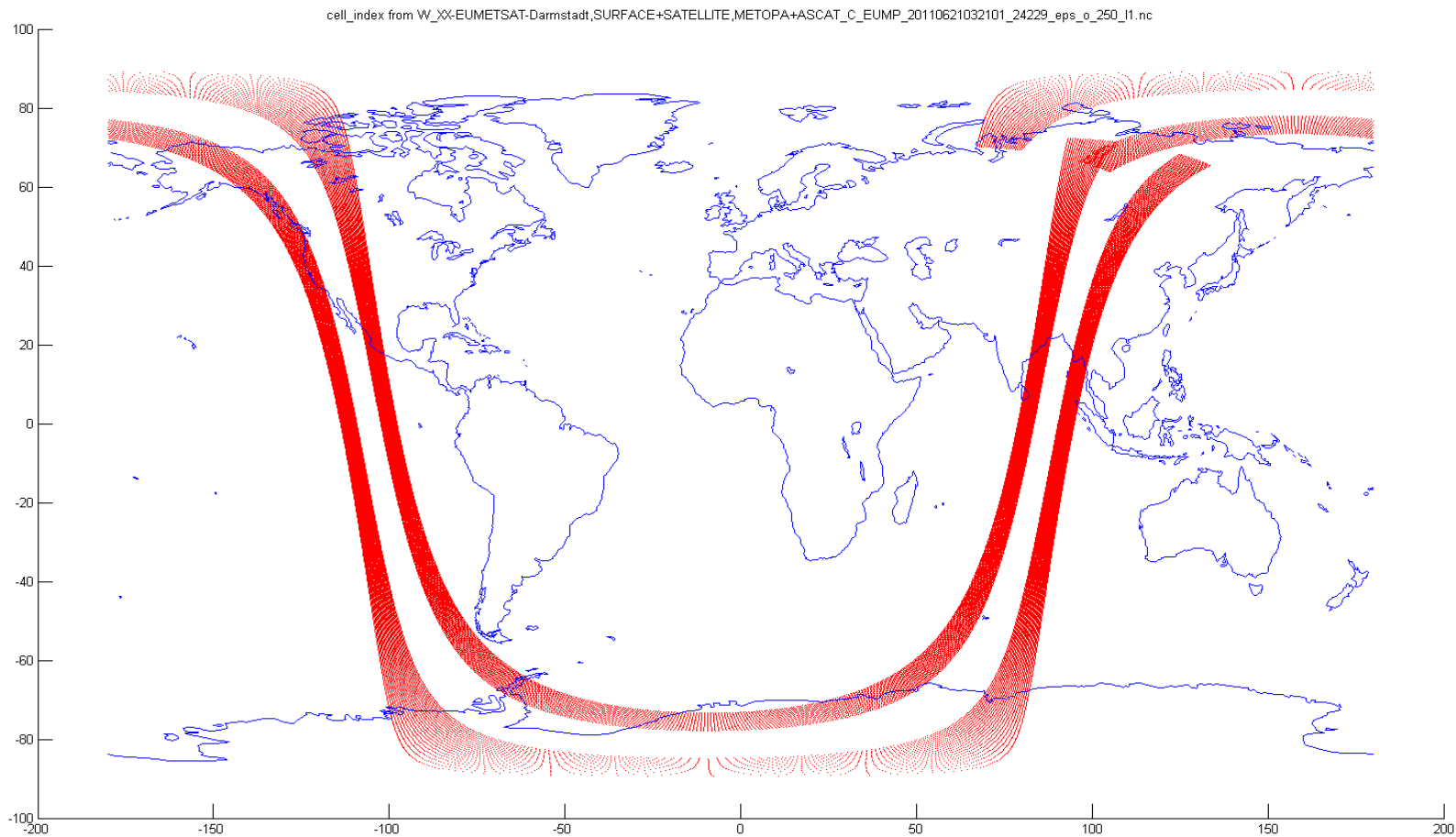
NetCDF Visualisation Tools: IDL - ASCAT L1b Plot



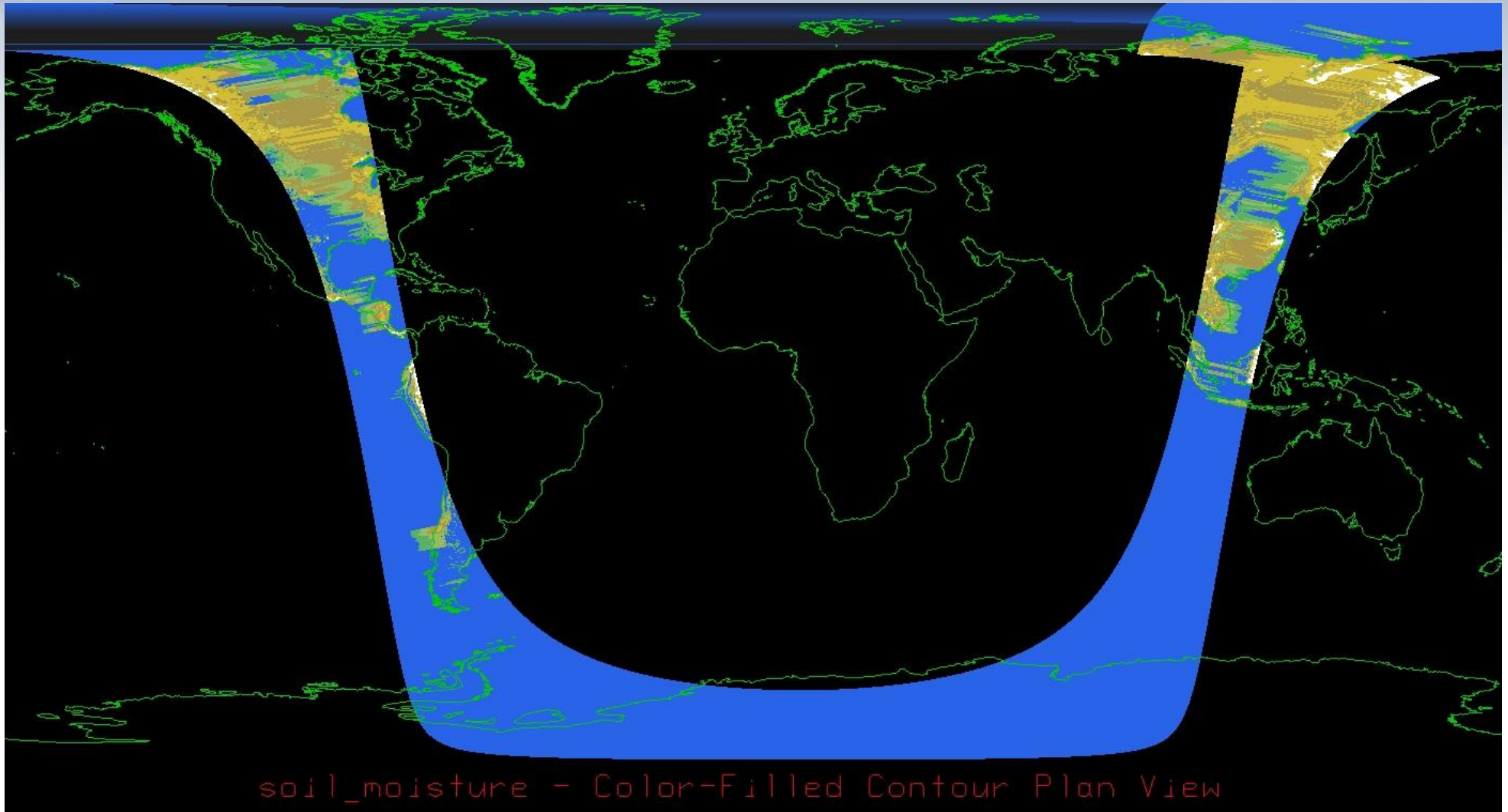
NetCDF Visualisation Tools : IDV - ASCAT L1b Plot



NetCDF Visualisation Tools : MatLab - ASCAT L1b Plot



NetCDF Visualisation Tools : IDV - ASCAT Soil Moisture L2 Plot



Steps for Validating the NetCDF Formats for Operations

The **goal** of the validation is to ensure the **quality** and the **usability** of the formats created:

- Quality is validated by ensuring each NetCDF format contains all the data fields with their contents from the Native product needed by the general user community. This validation is achieved partially by the instrument experts but mainly from YOUR feedback, the users.
- Usability is validated through evaluating the ease of working with the formats. NetCDF is inherently usable as there are many tools available to access the its contents. For data that can be visualised, the array data in the NetCDF files should provide meaningful plots in existing NetCDF visualisation tools with no or minimum development efforts.
- Data quality/accuracy is not within the scope of this pilot. Improvements in this area are made through **Reprocessing** to recreate the native products. Improvements in the native product will filter into the NetCDF formats when they are re-ordered once **Reprocessing** is completed.

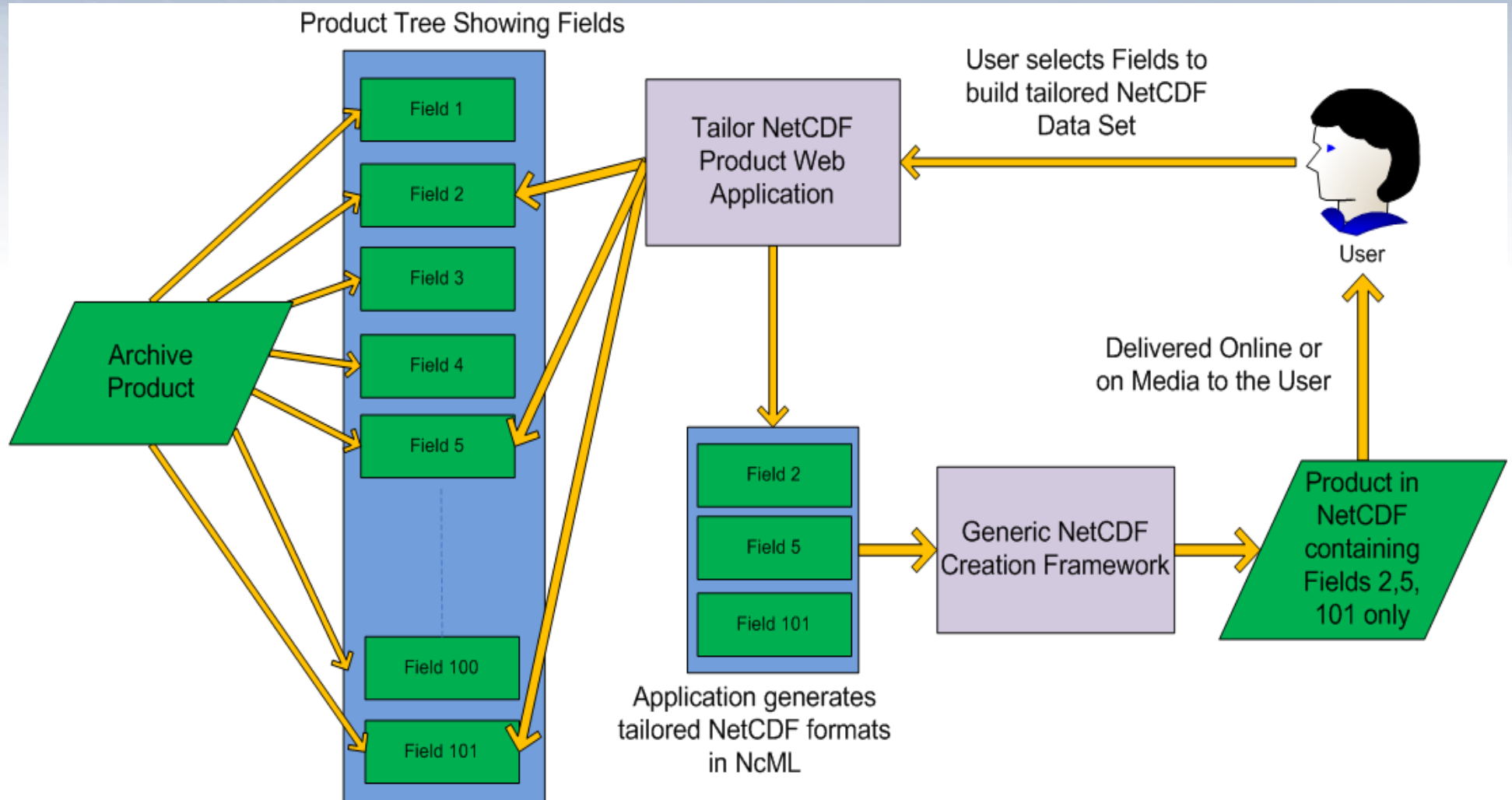
What have we learnt from the Pilot ?

- Instrument Experts with support from software developers are needed to create the product to NetCDF-component mappings.
- Software developers are required to create the XML configuration files that instruct the framework on how to create the NetCDF data sets. Climate and Forecast (CF) Conventions have been used where possible. New meta-data definitions to the CF conventions are to be discussed and proposed to the CF Conventions Committee.
- Software developers are required to code the JAVA classes to read the data from the products to populate the NetCDF components. A utility tool has been developed to create the JAVA class method stubs to aid implementation.
- The time needed to create a product's NetCDF format can now be estimated.

Implementing NetCDF as the Data Centre Archive's Common Delivery Format – Roadmap

- For the ASCAT NetCDF formats implemented, validation is expected to be completed so that the Formats are available for order from the Data Centre Archive by the 1st Quarter 2012.
- The framework prototype is provided as input to the full implementation process. Whether this is performed as a EUMETSAT in-house development or external contract is TBD.
- The Data Centre Archive tentatively schedules to offer LEO, GEO & (SAF) NetCDF formats for order by 2013. Users are invited to provide us with a **Wish List** and **Priorities** for products development.
- If there is a need, a further upgrade is to investigate the use of the framework for offering tailored NetCDF formats to the user community.

Possible Enhancement: Tailored NetCDF Formats



Useful Links

GSICS Data and Products Server: <http://gsics.eumetsat.int/>

GSICS User Notification Service: <http://eepurl.com/dB526>

Administrators have been asked to provide a more meaningful URL



Data Centre Archive Online Ordering Application: <http://archive.eumetsat.int/>

Product Navigator: <http://navigator.eumetsat.int/>

Earth Observation (EO) Portal: <http://eoportal.eumetsat.int/>

EUMETSAT Website: <http://www.eumetsat.int/>

We welcome your feedback on the Data Centre Service

Please provide it

- During the Conference → personally or at EUMETSAT booth
- Any other time → EUMETSAT helpdesk: ops@eumetsat.int